Selection of Optimal Lag Length in Cointegrated VAR Models with Weak Form of Common Cyclical Features

Carlos Enrique Carrasco Gutierrez and Reinaldo Castro Souza and Osmani Teixeira de Carvalho Guillén

2009

Selection of Optimal Lag Length in Cointegrated VAR Models with Weak Form of Common Cyclical Features*

Carlos Enrique Carrasco Gutierrez**
Reinaldo Castro Souza***
Osmani Teixeira de Carvalho Guillén****

Abstract

An important aspect of empirical research based on the vector autoregressive (VAR) model is the choice of the lag order, since all inferences in this model depend on the correct model specification. There have been many studies on how to select the lag order of a nonstationary VAR model subject to cointegration restrictions. In this paper, we consider an additional weak-form (WF) restriction of common cyclical features in the model to analyze the appropriate way to select the correct lag order. We use two methodologies: the traditional information criteria (AIC, HQ and SC) and an alternative criterion ($IC(p, s)$) that selects the lag order p and the rank structure s due to the WF restriction. We use a Monte Carlo simulation in the analysis. The results indicate that the cost of ignoring additional WF restrictions in vector autoregressive modeling can be high, especially when the SC criterion is used.

Keywords: Cointegration, Common Cyclical Features, Reduced Rank Model, Information Criteria.

JEL Codes: C32, C53.

---

*Submitted in February 2009. Revised in August 2009. We are grateful for the comments and suggestions by João Victor Issler, Wagner Gaglianone, Ricardo Cavalcanti, Luiz Renato Lima and many seminar participants at the 2006 Brazilian Econometric Meeting and 2007 European Econometric Meeting in Hungary. We extend special thanks to Alain Hecq for solving doubts and providing helpful comments. We are responsible for any remaining errors in this paper. Carlos Enrique also acknowledges financial support from Getulio Vargas Foundation and CAPES (Coordination for the Improvement of Higher Level Personnel).

**FUCAPE Business School, Vitória, Brazil. E-mail: carlos.gutierrez@fucape.br
***DEE-PUC-RJ, Brazil. E-mail: reinaldo@ele.puc-rio.br
****Banco Central do Brasil e Ibmec-RJ. E-mail: osmani.guillen@bcb.gov.br

1. Introduction

In the modeling of economic and financial time series, the vector autoregressive (VAR) model has become the standard linear model used in empirical works. An important aspect of empirical research on the specification of VAR models is determination of the lag order of the autoregressive lag polynomial, since all inferences in the VAR model depend on the correct model specification. Several works have demonstrated the effect of lag length selection. Lütkepohl (1993) indicated that selecting a higher order lag length than the true one causes an increase in the mean square forecast errors of the VAR and that underfitting the lag length often generates autocorrelated errors. Braun and Mittnik (1993) showed that impulse response functions and variance decompositions are inconsistently derived from the estimated VAR when the lag length differs from the true length. When cointegration restrictions are considered in the model, the effect of lag length selection on the cointegration tests has been demonstrated. For example, Johansen (1991) and Gonzalo (1994) pointed out that VAR order selection can affect proper inference about cointegrating vectors and rank.

Recently, empirical works have considered other kinds of restrictions in the VAR model (e.g., Engle and Issler (1995); Caporale (1997); Mamingi and Iyare (2003)). Engle and Kozicki (1993) showed that VAR models can have other types of restrictions, called common cyclical features, which are restrictions on the short-run dynamics. These restrictions are defined in the same way as cointegration restrictions, but while cointegration refers to relations among variables in the long run, common cyclical restrictions refer to relations in the short run. Vahid and Engle (1993) proposed the serial correlation common feature (SCCF) as a measure of common cyclical features. SCCF restrictions might be imposed in a covariance stationary VAR model or in a cointegrated VAR model. The concept of serial correlation common features appears to be useful. It means that stationary time series move together in a way such that there are linear combinations of these variables which yield white noise processes and that their impulse response functions are collinear. In several practical applications the existence of short-run comovements between stationary time series (e.g., between first-differenced cointegrated I(1)) has been analyzed. For instance, Engle and Issler (1995) found common cycle comovement in U.S. sectoral output data; Hecq (2002) and Engle and Issler (1993) found common cycles in Latin American countries; and Carrasco and Gomes (2009) found common international cycles in GNP data for Mercosur countries.

When short-run restrictions are imposed in cointegrated VAR models, it is possible to define a weak version of SCCF restrictions. Hecq et al. (2006) defined a weak version of SCCF restrictions, which they denominated weak-form (WF) common cyclical restrictions. A fundamental difference between SCCF and WF restrictions is the form in which each one imposes restrictions on the representation of the vector error correction model (VECM).[1] When SCCF are imposed, all matrices of a VECM have rank less than the number of variables analyzed. On the other hand, with WF restrictions, all matrices except the long-run matrix have rank less than the number of variables being analyzed. Hence, WF restrictions impose less constraint on VECM parameters. Some advantages emerge when WF restrictions are considered. First, due to the fact that the weak-form common cyclical method does not impose constraints on the cointegration space, the rank of common cyclical features is not limited by the choice of cointegrating rank.

The literature has shown how to select an adequate lag order of a covariance stationary VAR model and an adequate lag order of a VAR model subject to cointegration restrictions. Among the classical procedures are information criteria, such as Akaike (AIC), Schwarz (SC) and Hannan-Quinn (HQ) (Lütkepohl, 1993). Kilian (2001) studied the performance of traditional AIC, SC and HQ criteria of a covariance stationary VAR model. Vahid and Issler (2002) analyzed the standard information criteria in a covariance stationary VAR model subject to SCCF restriction and, more recently, Guillén et al. (2005) studied the standard information criteria in VAR models with cointegration and SCCF restrictions. However, when cointegrated VAR models contain an additional weak form of common cyclical features, there are no reported works on how to appropriately determine the VAR model order.

The objective of this paper is to investigate the performance of information criteria in selecting the lag order of a VAR model when the data are generated from a true VAR with cointegration and WF restrictions, referred to as the correct model. We carry out the following two procedures:

a) the use of standard criteria, as proposed by Vahid and Engle (1993), referred to here as IC(p), and

b) the use of an alternative model selection criterion (see Vahid and Issler (2002) and Hecq et al. (2006)), which consists in simultaneously selecting the lag order p and the number of weak forms of common cyclical features, s, which is referred to as IC(p, s).[2]

The most relevant results can be summarized as follows. The information criterion that selects the pair (p, s) performs better than the model chosen by conventional criteria, especially the AIC(p, s) criterion. The cost of ignoring additional WF restrictions in vector autoregressive modeling can be high, particularly when the SC(p) criterion is used.

The rest of this paper is organized as follows. Section 2 presents the econometric model. Section 3 discusses the information criteria. Section 4 shows a Monte

[1]When a VAR model has cointegration restriction it can be represented as a VECM. This representation is also known as Granger representation theorem (Engle and Granger, 1987).

[2]This is quite recent in the literature (see, Hecq (2006)).

Carlo simulation and Section 5 presents the results. Finally, Section 6 contains our conclusions.

2. The Econometric Model

We show the VAR model with short-run and long-run restrictions. First, we consider a Gaussian vector autoregression of finite order p, called VAR(p), such that:

$$y_t = \sum_{i=1}^{p} A_i y_{t-i} + \varepsilon_t \tag{1}$$

where, y_t is a vector of n first-order integrated series, $I(1)$, A_i, $i = 1, \ldots, p$ are matrices of dimension $n \times n$, $\varepsilon_t \sim Normal\,(0, \Omega)$, $E\,(\varepsilon_t) = 0$ and $E\,(\varepsilon_t \varepsilon'_\tau)$ $=\{\Omega$, if $t = \tau$ and $0_{n\times n}$, if $t \neq \tau$, where Ω is nonsingular$\}$. The model (1) can be written equivalently as; $\Pi\,(L)\, y_t = \varepsilon_t$ where L represents the lag operator and $\Pi\,(L) = I_n - \sum_{i=1}^{p} A_i L^i$ such that when $L = 1$, $\Pi\,(1) = I_n - \sum_{i=1}^{p} A_i$. If cointegration is considered in (1) the $(n \times n)$ matrix $\Pi\,(\cdot)$ satisfies two conditions: a) rank $(\Pi\,(1)) = r$, $0 < r < n$, such that $\Pi\,(1)$ can be expressed as $\Pi\,(1) = -\alpha\beta'$, where α and β are $(n \times r)$ matrices with full column rank, r; and b) the characteristic equation $|\Pi\,(L)| = 0$ has $n - r$ roots equal to 1 and all others are outside the unit circle. These assumptions imply that y_t is cointegrated of order $(1, 1)$. The elements of α are the adjustment coefficients and the columns of β span the space of cointegration vectors. We can represent a VAR model as a vector error correction model (VECM). By decomposing the polynomial matrix $\Pi\,(L) = \Pi\,(1)\, L + \Pi^*\,(L)\,\Delta$, where $\Delta \equiv (1 - L)$ is the difference operator, a VECM is obtained:

$$\Delta y_t = \alpha\beta' y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \varepsilon_t \tag{2}$$

where: $\alpha\beta' = -\Pi\,(1)$, $\Gamma_j = -\sum_{k=j+1}^{p} A_k$ for $j = 1, ..., p-1$ and $\Gamma_0 = I_n$. The VAR(p) model can include additional short-horizon restrictions as shown by Vahid and Engle (1993). We consider an interesting WF restriction (as defined by Hecq et al. (2006)) that does not impose constraints on long-run relations.

Definition 1 *The weak form (WF) holds in (2) if, in addition to cointegration restriction, there exists an $(n \times s)$ matrix $\tilde{\beta}$ of rank s, whose columns span the cofeature space, such that $\tilde{\beta}'(\Delta y_t - \alpha\beta y_{t-1}) = \tilde{\beta}'\varepsilon_t$, where $\tilde{\beta}'\varepsilon_t$ is an s-dimensional vector that constitutes an innovation process with respect to information prior to period t, given by $\{y_{t-1},\ y_{t-2}, ..., y_1\}$.*

Equivalent to Definition 1, we consider WF restrictions in the VECM if there exists a cofeature matrix $\tilde{\beta}$ that satisfies the following assumption:

Assumption 1: $\tilde{\beta}' \, \Gamma_j = 0_{s\times n}$ *for* $\quad j = 1, ..., p-1.$

Therefore, this is a naturally weaker alternative assumption which implies that the common cyclical part is reduced to white noise by taking a linear combination of the variables in the first differences adjusted for long-run effects. Imposing WF restrictions is convenient because it allows for the study of both cointegration and common cyclical feature without the constraint[3] $r + s \leq n$.

We can rewrite the VECM with WF restrictions as a model of reduced-rank structure. In (2) let $X_{t-1} = [\Delta y'_{t-1}, ...\Delta y'_{t-p+1}]'$ and $\Phi = [\Gamma_1, ..., \Gamma_{p-1}]$. Therefore, we obtain:

$$\Delta y_t = \alpha\beta y_{t-1} + \Phi X_{t-1} + \varepsilon_t \tag{3}$$

If Assumption (1) holds, then matrices $\Gamma_i, i = 1, ..., p$ are all of rank $(n-s)$ and we can write $\Phi = \tilde{\beta}_\perp \Psi = \tilde{\beta}_\perp[\Psi_1, ..., \Psi_{p-1}]$, where, $\tilde{\beta}_\perp$ is $n \times (n-s)$ full column rank matrix, Ψ has dimension $(n-s)\times n(p-1)$, and the matrices $\Psi_i, i = 1, ..., p-1$ all have rank $(n-s) \times n$. Hence, given Assumption (1), there exists $\tilde{\beta}$ of $n \times s$ such that $\tilde{\beta}'\tilde{\beta}_\perp = 0$. That is, $\tilde{\beta}_\perp$ $n \times (n-s)$ is a full column rank orthogonal to the complement of $\tilde{\beta}$ with $rank(\tilde{\beta}, \tilde{\beta}_\perp) = n$. Rewriting model (3) we have:

$$\begin{aligned} \Delta y_t &= \alpha\beta y_{t-1} + \tilde{\beta}_\perp \left(\Psi_1, \Psi_2, ..., \Psi_{p-1}\right) X_{t-1} + \varepsilon_t && (4)\\ &= \alpha\beta y_{t-1} + \tilde{\beta}_\perp \Psi X_{t-1} + \varepsilon_t && (5) \end{aligned}$$

Estimation of (5) is carried out via the switching algorithms (see Centoni et al. (2007); Hecq (2006)) that use the procedure in estimating reduced-rank regression models as suggested by Anderson (1951). There is a formal connection between a reduced-rank regression and the canonical analysis, as noted by Izenman (1975), Box and Tiao (1977), Tso (1981) and Velu et al. (1986). When all the matrix coefficients of the multivariate regression are full rank, they can be estimated by the usual least squares or maximum likelihood procedures. But when the matrix coefficients are of reduced rank they have to be estimated using the reduced-rank regression models of Anderson (1951). The use of canonical analysis may be regarded as a special case of reduced-rank regression. More specifically, the maximum likelihood estimation of the parameters of the reduced-rank regression model may solve a problem of canonical analysis.[4] Therefore, we can use the expression $CanCorr\{X_t, Z_t|X_{t-1}\}$ which denotes the partial canonical correlations between

[3] Since the SCCF also imposes constraints on the long-run matrix $\alpha\beta' = -\Pi(1)$, which has dimension n, the cointegration restrictions, r, and SCCF restrictions, s, must satisfy $r + s \leq n$.

[4] This estimation is referred to as full information maximum likelihood – FIML. (see Johansen (1995).

X_t and Z_t: both sets are concentrated out of the effect of X_{t-1} allowing us to obtain canonical correlation (see Johansen (1995)), represented by the eigenvalues $\hat{\lambda}_1 > \hat{\lambda}_2 > \hat{\lambda}_3 .... > \hat{\lambda}_n$. The Johansen test statistic is based on canonical correlation. In model (2) we can use the expression $CanCorr\{\Delta y_t, y_{t-1}|X_{t-1}\}$ where $X_{t-1} = [\Delta y'_{t-1}, ...\Delta y'_{t-p+1}]'$, which summarizes the reduced-rank regression procedure used in Johansen approach. This means that we extract the canonical correlations between Δy_t and y_{t-1}: both sets are concentrated out of the effect of lags of X_{t-1}. In order to test for the significance of the r largest eigenvalues, we can rely on Johansen's trace statistic (6):

$$\xi_r = -T \sum_{i=r+1}^{n} Ln\,(1 - \hat{\lambda}_i^2) \qquad i = 1, ..., n \tag{6}$$

where the eigenvalues $0 < \hat{\lambda}_n < ... < \hat{\lambda}_1$ are the solution to : $|\lambda m_{11} - m_{10}^{-1} m_{00} m_{01}| = 0$, where m_{ij}, i, $j = 0.1$, are the second-moment matrices: $m_{00} = \frac{1}{T}\sum_{t=1}^{T} \tilde{u}_{0t}\tilde{u}'_{0t}$, $m_{10} = \frac{1}{T}\sum_{t=1}^{T} \tilde{u}_{1t}\tilde{u}'_{0t}$, $m_{01} = \frac{1}{T}\sum_{t=1}^{T} \tilde{u}_{0t}\tilde{u}'_{1t}$, $m_{11} = \frac{1}{T}\sum_{t=1}^{T} \tilde{u}_{1t}\tilde{u}'_{1t}$ of the residuals $\tilde{u}_{0t}$ and $\tilde{u}_{1t}$ obtained in the multivariate least squares regressions $\Delta\ y_t = (\Delta y_{t-1},...,\Delta y_{t-p+1})\ +u_{0t}$ and $y_{t-1} = (\Delta y_{t-1}, ...\ \Delta y_{t-p+1})\ +u_{1t}$ respectively (see, Hecq (2006); Johansen (1995)). The result of Johansen test is a superconsistent estimation of β. Moreover, we could also use a canonical correlation approach to determine the rank of the common feature space due to WF restrictions. This is a test for the existence of cofeatures in the form of linear combinations of the variables in first differences, corrected for long-run effects which are white noise (i.e., $\tilde{\beta}'(\Delta y_t - \alpha\beta y_{t-1}) = \tilde{\beta}'\varepsilon_t$ where $\tilde{\beta}'\varepsilon_t$ is a white noise). We use canonical analysis is this work to estimate and select the lag rank of VAR models, as shown in the subsequent sections.

3. Model Selection Criteria

In model selection we use two procedures to identify the VAR model order: the standard selection criteria, IC(p), and the modified information criteria, IC(p, s), a novelty in the literature, which consists in identifying p and s simultaneously.

The model estimation following the standard selection criteria, IC(p), originally used by Vahid and Engle (1993), entails the following steps:

1. Estimate p using standard information criteria: Akaike (AIC), Schwarz (SC) and Hanna-Quinn (HQ). We chose the lag length of the VAR in levels that minimizes the information criteria.

2. Using the lag length chosen in the previous step, find the number of cointegration vectors, r, using Johansen cointegration test.[5]

[5]Cointegration rank and vectors are estimated using the FIML, as shown in Johansen (1991).

3. Conditional on the results of the cointegration analysis, estimate a final VECM and then calculate the multi-step ahead forecast.

The above procedure is followed when there is evidence of cointegration restrictions. We check the performance of IC(p) when WF restrictions are imposed on the true model. Additionally, we check the performance of IC(p, s) alternative selection criteria. Vahid and Issler (2002) analyzed a covariance stationary VAR model with SCCF restrictions. They showed that the use of IC(p, s) performs better than IC(p) in VAR model lag order selection. In the present paper we analyze the cointegrated VAR model with WF restrictions in order to analyze the performance of IC(p) and IC(p, s) for model selection. The question investigated is: Does IC(p, s) perform better than IC(p)? This is an important question we aim to answer in this paper.

The procedure to choose the lag order and the rank of the structure of short-run restrictions is carried out by minimizing the following modified information criteria (see Vahid and Issler (2002); Hecq (2006)).

$$AIC\,(p,s) = \sum_{i=n-s+1}^{T} \ln(1-\lambda_i^2(p)) + \frac{2}{T} \times N \tag{7}$$

$$HQ(p,s) = \sum_{i=n-s+1}^{T} \ln(1-\lambda_i^2(p)) + \frac{2\ln(\ln T)}{T} \times N \tag{8}$$

$$SC(p,s) = \sum_{i=n-s+1}^{T} \ln(1-\lambda_i^2(p)) + \frac{\ln T}{T} \times N \tag{9}$$

$$N = [n \times (n \times (p-1)) + n \times r] - [s \times (n \times (p-1) + (n-s))]$$

where n is the number of variables in model (2) and N is a number of parameters. N is obtained by subtracting the total number of mean parameters in the VECM (i.e., $n^2 \times (p-1) + nr$), for given r and p, from the number of restrictions the common dynamics imposes from $s \times (n \times (p-1)) - s \times (n-s)$. The eigenvalues λ_i are calculated for each p. In order to calculate the pair (p, s) we assume that no restriction exists, that is, $r = n$ (see Hecq (2006)). We fix p in model (3) and then find λ_i $i = 1, 2...n$ by computing the $cancorr(\Delta y_t, X_{t-1} \mid y_{t-1})$. This procedure is followed for every p and in the end we choose the p and s that minimize the IC(p, s). After selecting the pair (p, s) we can test the cointegration relation using the Johansen procedure. Finally, we estimate the model using the switching algorithms as shown in the next section. Notice that in this simultaneous selection, testing the cointegration relation is the last procedure followed, so we are inverting the hierarchical procedure followed by Vahid and Engle (1993) where the first step is to select the number of cointegration relations. This may be an advantage, especially when r is overestimated. Few works have analyzed the order of VAR models considering modified IC(p, s). As mentioned, Vahid and Issler (2002) suggested the use of IC(p, s) to simultaneously choose the order p and a number of reduced-rank structure s in a covariance stationary VAR model subject to SCCF restrictions. However, no work has analyzed the order of the VAR model with cointegration and WF restrictions using a modified criterion, which is exactly the contribution of this paper.

To estimate the VAR model, considering cointegration and WF restrictions, we use the switching algorithms model as considered by Hecq (2006). Consider the VECM given by:

$$\Delta y_t = \alpha\beta' y_{t-1} + \tilde{\beta}_\perp \Psi X_{t-1} + \varepsilon_t \tag{10}$$

A full description of switching algorithms is presented below in four steps:

*Step*1 : Estimation of the cointegration vectors β.

Using the optimal pair $(\bar{p}, \bar{s})$ chosen by the information criteria (7), (8) or (9), we estimate β (and so its rank, $r = \bar{r}$) using Johansen cointegration test.

*Step*2 : Estimation of $\tilde{\beta}_\perp$ and Ψ.

Taking the estimate of β in step one, we proceed to estimate $\tilde{\beta}_\perp$ and Ψ. Hence, we run a regression of Δy_t and of X_{t-1}on $\hat{\beta}' y_{t-1}$. We label the residuals u_0 and u_1, respectively. Therefore, we obtain a reduced-rank regression:

$$u_{0t} = \tilde{\beta}_\perp \Psi u_{1t} + \varepsilon_t \tag{11}$$

where Ψ can be written as $\Psi = \left(C_1, ..., C_{(\bar{p}-1)}\right)$ of $(n - \bar{s}) \times n(\bar{p} - 1)$ and $\tilde{\beta}_\perp$ of $n \times (n - \bar{s})$. We estimate (11) by FIML. Thus, we can obtain $\tilde{\beta}_\perp$ and $\hat{\Psi}$.

*Step*3 : Estimation of the maximum likelihood (ML) function.

Given the parameters estimated in steps 1 and 2 we use a recursive algorithm to estimate the maximum likelihood (ML) function. We calculate the eigenvalues associated with $\hat{\Psi}$, $\hat{\lambda}_i^2$ $i = 1, ..., \bar{s}$ and the matrix of residuals $\sum_{\bar{r},\ s=\bar{s}}^{\max}$. Hence, we compute the ML function:

$$L^0_{\max,\ \bar{r}<n,\ s=\bar{s}} = -\frac{T}{2}\left[\ln\left|\sum_{\bar{r}<n,\ s=\bar{s}}^{\max}\right| - \sum_{i=1}^{\bar{s}} \ln\left(1 - \hat{\lambda}_i^2\right)\right] \tag{12}$$

If $\bar{r} = n$, instead of (12) we use the derived log-likelihood: $L_{\max,\ r=n,\ s=\bar{s}} = -\frac{T}{2} \ln \left| \sum_{\bar{r}=n,\ s=\bar{s}}^{\max} \right|$. The determinant of the covariance matrix for $\bar{r} = n$ cointegration vector is calculated by

$$\ln \left| \sum_{\bar{r}=n,\ s=\bar{s}}^{\max} \right| = \ln \left| m_{00} - m_{01} m_{11}^{-1} m_{10} \right| - \sum_{i=1}^{\bar{s}} \ln \left(1 - \hat{\lambda}_i^2 \right) \tag{13}$$

where m_{ij} refers to cross moment matrices obtained in multivariate least squares regressions from Δy_t and X_{t-1} on y_{t-1}. In this case, estimation does not entail using an iterative algorithm yet, because the cointegrating space spans R^n.

*Step*4 : Reestimation of β.

We reestimate β to obtain a more appropriate value for the parameters. In order to reestimate β we compute the $CanCorr \left[\Delta y_t, y_{t-1} \mid \hat{\Psi} X_{t-1} \right]$ and thus using the new $\hat{\beta}$ we can repeat step 2 to reestimate $\tilde{\beta}_\perp$ and Ψ. Then, we calculate the new value of the ML function in step 3. Hence, we obtain $L^1_{\max,\ r=\bar{r},\ s=\bar{s}}$ to calculate $\Delta L = (L^1_{\max,\ r=\bar{r},\ s=\bar{s}}$ - $L^0_{\max,\ r=\bar{r},\ s=\bar{s}})$.

We repeat steps 1 to 4 to choose $\tilde{\beta}_\perp$and Ψ until convergence is reached (i.e., $\Delta L < 10^{-7}$). In the end, the optimal parameters $\bar{p}$, $\bar{r}$ and $\bar{s}$ are obtained and they can be used to estimate and forecast a VECM with WF restrictions.

4. Monte Carlo Design

The simple real business cycle models and also the simplest closed economy monetary dynamic stochastic general equilibrium models are three-dimensional. Consumption, saving and output and prices, output and money are notable examples. Motivated by these applications and according to the previous paper of Vahid and Issler (2002), we construct a Monte Carlo experiment in a three-dimensional environment. Therefore, the data generating processes considering a VAR model with three variables, one cointegration vector, and two cofeature vectors (i.e., $n = 3$, $r = 1$ and $s = 2$, respectively). β and $\tilde{\beta}$ satisfy:

$$\beta = \begin{bmatrix} 1.0 \\ 0.2 \\ -1.0 \end{bmatrix}, \tilde{\beta} = \begin{bmatrix} 1.0 & 0.1 \\ 0.0 & 1.0 \\ 0.5 & -0.5 \end{bmatrix}$$

$$\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1.0 & 0.6 & 0.6 \\ 0.6 & 1.0 & 0.6 \\ 0.6 & 0.6 & 1.0 \end{bmatrix} \right)$$

Consider the VAR(3) model: $y_t = A_1 y_{t-1} + A_2 y_{t-2} + A_3 y_{t-3} + \varepsilon_t$. The VECM representation as a function of the VAR level parameters can be written as:

$$\Delta y_t = (A_1 + A_2 + A_3 - I_3) y_{t-1} - (A_2 + A_3)\Delta y_{t-1} - A_3 \Delta y_{t-2} + \varepsilon_t \tag{14}$$

The VAR coefficients must simultaneously comply with the restrictions:

a) the cointegration restrictions: $\alpha\beta' = (A_1 + A_2 + A_3 - I_3)$;

b) WF restrictions: $\tilde{\beta}' A_3 = 0$ (iii) $\tilde{\beta}'(A_2 + A_3) = 0$ and c) the covariance stationary condition.

Considering the cointegration restrictions, we can rewrite (14) as the following VAR(1):

$$\xi_t = F\ \xi_{t-1} + v_t \tag{15}$$

$$\xi_t = \begin{bmatrix} \triangle y_t \\ \triangle y_{t-1} \\ \beta' y_t \end{bmatrix},\ F = \begin{bmatrix} -(A_2 + A_3) & -A_3 & \alpha \\ I_3 & 0 & 0 \\ -\beta(A_2 + A_3) & -\beta' A_3 & \beta'\alpha + 1 \end{bmatrix} \text{ and } v_t = \begin{bmatrix} \varepsilon_t \\ 0 \\ \beta'\varepsilon_t \end{bmatrix}$$

Thus, Equation (15) will be covariance stationary if all eigenvalues of matrix F lie inside the unit circle. An initial idea to design the Monte Carlo experiment can consist in constructing the companion matrix (F) and verifying whether the eigenvalues of the companion matrix all lie inside the unit circle. This can be carried out by selecting their values from a uniform distribution, and then verifying whether or not the eigenvalues of the companion matrix all lie inside the unit circle. However, this strategy could lead to a wide spectrum of search for adequate values for the companion matrix. Hence, we follow an alternative procedure. We propose an analytical solution to generate a covariance stationary VAR, based on the choice of the eigenvalues, and then on the generation of the respective companion matrix. In the Appendix, we present a detailed discussion of the final choice of these free parameters, including analytical solutions. In our simulation, we constructed 100 data generating processes and for each of these we generated 1,000 samples containing 1,000 observations. To reduce the impact of the initial values, we considered only the last 100 and 200 observations. All the experiments were conducted in the MatLab environment.

5. Results

Figure 1 shows one example of the three-dimensional VAR model with cointegration and WF restrictions for 100 and 200 observations.

The values in Table A.1 represent the percentage of time the model selection criterion, IC(p), takes to choose the cell corresponding to the lag and number of

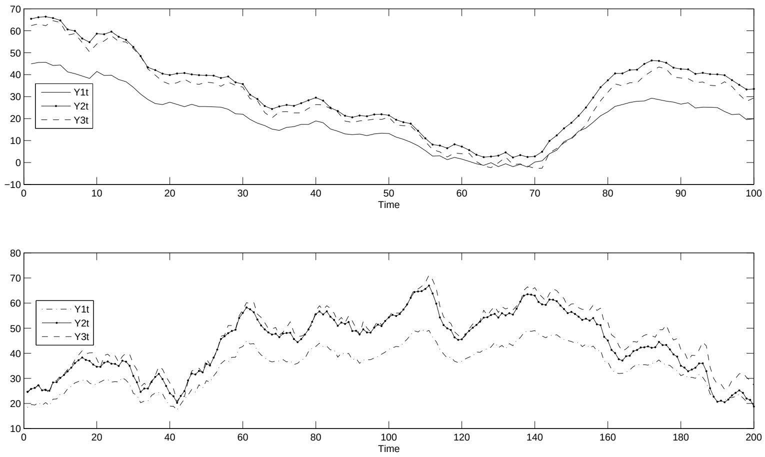

Figure 1
One example of a VAR(3) model with $n = 3$, $r = 1$ and $s = 2$ for 100 and 200 observations

cointegration vectors in 100,000 runs. The true lag-cointegrating vectors are identified by boldface numbers and the selected lag-cointegration vectors often chosen by the criterion are underlined. In Table A.1, the results show that, in general, the AIC most often chooses the correct lag length for 100 and 200 observations. For example, for 100 observations, the AIC, HQ and SC criteria chose the true lag, p, 54.08%, 35.62% and 17.48% of the times, respectively. Note that all three criteria chose the correct rank of cointegration ($r = 1$). When 200 observations were considered, the correct lag length was chosen 74.72%, 57.75% and 35.28% of the times for AIC, HQ and SC, respectively. Again, all three criteria selected the true cointegrated rank $r = 1$. Table A.2 contains the percentage the alternative model selection criterion, IC(p, s), has in choosing that cell, corresponding to the lag rank and number of cointegrating vectors in 100,000 runs. The true lag rank cointegration vectors are identified by boldface numbers and the best lag rank combination often chosen by each criterion are underlined. In Table A.2, the results show that, in general, the AIC(p, s) criterion more frequently chooses the lag rank for 100 and 200 observations. For instance, for 100 observations, the AIC(p, s), HQ(p, s) and SC(p, s) criteria more often choose the true pair $(p, s) = (3, 1)$, 56.34%, 40.85% and 25.2% of the times, respectively. For 200 observations, the AIC(p, s), HQ(p, s) and SC(p, s) criteria more frequently choose the true pair $(p, s) = (3, 1)$, 77.06%, 62.58% and 45.03% of the times, respectively. Note that all three criteria more often choose the correct rank of cointegration ($r = 1$) in both samples. What happens when the weak-form common cyclical restrictions are ignored? Tables A.1 and A.2 also show the relative performance of IC(p, s) vis-à-vis IC(p). For instance, for $T = 100$ the SC(p, s) has a success rate of 25.2% in selecting the true $p = 3$, while the SC(p) only has a success rate of 17.48%. This represents a gain of more than 44%. For $T = 200$, the gains are more than 27%. For $T = 100$ the HQ(p, s) selects the true p =3 with a 40.85% accuracy while the HQ(p) only has a success rate of 35.62%. This represents a gain of 14%. For $T = 200$, the gains are more than 8%. For $T = 100$ the AIC(p, s) has a success rate of 56.34% in choosing the true $p = 3$, in comparison with a rate of 54.08% for the AIC(p), a gain of more than 4%. For T=200, the gains are more than 3%. Thus, it appears that when using the AIC(p,s) criteria the cost of ignoring the weak-form common cyclical restriction is low.

The most relevant results can be summarized as follows:

− All criteria (AIC, HQ and SC) choose the correct parameters more often when using IC(p, s) vis-à-vis IC(p).

− There is a cost of ignoring additional weak-form common cyclical restrictions in the model especially when the SC(p) criterion is used. In general, the standard Schwarz, SC(p), or Hannan-Quinn, HQ(p), selection criteria should not be used for this purpose in small samples due to the tendency to identify an underparameterized model.

− The AIC performs better in selecting the true model more frequently for both the IC(p, s) and the IC(p) criteria.

6. Conclusions

In this paper, we considered an additional weak-form restriction of common cyclical features in a cointegrated VAR model in order to analyze the appropriate way to choose the correct lag order. These additional WF restrictions are defined in the same way as cointegration restrictions. While cointegration refers to relations among variables in the long run, the common cyclical restrictions refer to relations in the short run. Two methodologies have been used for selecting lag length; the traditional information criterion, IC(p), and an alternative criterion (IC(p, s)) that selects the lag order p and the rank structure s due to the weak-form common cyclical restrictions.

The results indicate that the information criterion that selects the lag length and the rank order performs better than the model chosen by conventional criteria. When the WF restrictions are ignored there is a nontrivial cost in selecting the true model with standard information criteria. In general, the standard Schwarz or Hannan-Quinn selection criteria should not be used for this purpose in small samples, due to the tendency to identify an underparameterized model.

In applied work, when the VAR model contains WF and cointegration restrictions, we suggest the use of AIC(p, s) criteria to choose the lag rank, since it provides considerable gains in selecting the correct VAR model. Since no work in the literature has analyzed a VAR model with WF common cyclical restrictions, the results of this paper provide new insights and incentives to proceed with this kind of empirical work.

References

Anderson, T. W. (1951). Estimating linear restrictions on regression coefficients for multivariate normal distributions. *The Annals of Mathematical Statististics*, 22:327–351.

Box, G. E. & Tiao, G. C. (1977). A canonical analysis of multiple time series. *Biometrika*, 64:355–365.

Braun, P. A. & Mittnik, S. (1993). Misspecifications in vector autoregressions and their effects on impulse responses and variance decompositions. *Journal of Econometrics*, 59:319–41.

Caporale, G. M. (1997). Common features and output fluctuations in the United Kingdom. *Economic Modelling*, 14:1–9.

Carrasco, C. G. & Gomes, F. (2009). Evidence on common features and business cycle synchronization in Mercosur. *Brazilian Review of Econometrics*, 29:forthcoming.

Centoni, M., Cubadda, G., & Hecq, A. (2007). Common shocks, common dynamics and the international business cycle. *Economic Modelling*, 24:149–166.

Engle, R. & Granger, C. (1987). Cointegration and error correction: Representation, estimation and testing. *Econometrica*, 55:251–76.

Engle, R. & Issler, J. V. (1993). Common trends and common cycles in Latin America. *Revista Brasileira de Economia*, 47:149–176.

Engle, R. & Issler, J. V. (1995). Estimating common sectoral cycles. *Journal of Monetary Economics*, 35:83–113.

Engle, R. & Kozicki, S. (1993). Testing for common features. *Journal of Business and Economic Statistics*, 11:369–395.

Gonzalo, J. (1994). Five alternative methods of estimating long-run equilibrium. *Journal of Econometrics*, 60:203–233.

Guillén, O. T. C., Issler, J. V., & Athanasopoulos, G. (2005). Forecasting accuracy and estimation uncertainty using VAR models with short-and long-term economic restrictions: A Monte-Carlo study. Ensaios Econômicos EPGE 589.

Hecq, A. (2002). Common cycles and common trends in Latin America. *Medium Econometrische Toepassingen*, 10:20–25.

Hecq, A. (2006). Cointegration and common cyclical features in VAR models: Comparing small sample performances of the 2-step and interative approaches. Mimeo.

Hecq, A., Palm, F. C., & Urbain, J. P. (2006). Testing for common cyclical features in VAR models with cointegration. *Journal of Econometrics*, 132:117–141.

Izenman, A. J. (1975). Reduced rank regression for the multivariate linear model. *Journal of Multivariate Analysis*, 5:248–264.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoegressive models. *Econometrica*, 59:1551–1580.

Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press.

Kilian, L. (2001). Impulse response analysis in vector autoregressions with unknown lag order. *Journal of Forecasting*, 20:161–179.

Lütkepohl, H. (1993). *Introduction to Multiple Time Series Analysis*. Springer, New York.

Mamingi, N. & Iyare, S. O. (2003). Convergence and common features in international output: A case study of the economic community of West African States 1975-1997. *Asian African Journal of Economics and Econometrics*, 3:1–16.

Tso, M. K. (1981). Reduced rank regression and canonical analysis. *Journal of the Royal Statistical Society*, 43:183–189.

Vahid, F. & Engle, R. (1993). Common trends and common cycles. *Journal of Applied Econometrics*, 8:341–360.

Vahid, F. & Issler, J. V. (2002). The importance of common cyclical features in VAR analysis: A Monte Carlo study. *Journal of Econometrics*, 109:341–363.

Velu, R. P., Reinsel, G. C., & Wichern, D. W. (1986). Reduced rank models for multiple time series. *Biometrika*, 73:105–118.

Appendix A: Tables

Table A.1 Performance of the IC(p) information criterion in selecting lag order p

Frequency of lag(p) and cointegrating vectors (r) chosen by different criteria for the trivariate VAR model in levels when the true model has parameters: $p = 3$ and $r = 1$

		Number of observations=100 selected cointegrated vectors				Number of observations=200 selected cointegrated vectors			
		0	**1**	2	3	0	1	2	3
	Selected lag								
AIC(p)	1	0,000	0,996	0,359	0,031	0,000	0,095	0,016	0,003
	2	0,002	32,146	1,136	0,048	0,000	17,073	0,686	0,033
	3	2,792	**54,082**	0,902	0,041	0,012	**74,721**	1,488	0,108
	4	0,737	4,068	0,091	0,003	0,005	4,177	0,081	0,006
	5	0,392	0,987	0,031	0,000	0,013	0,828	0,020	0,000
	6	0,219	0,333	0,014	0,000	0,023	0,257	0,005	0,000
	7	0,166	0,173	0,006	0,000	0,039	0,133	0,002	0,000
	8	0,133	0,107	0,005	0,000	0,060	0,115	0,001	0,000
HQ(p)	1	0,000	3,884	1,915	0,165	0,000	1,098	0,243	0,021
	2	0,002	52,593	1,907	0,080	0,000	37,390	1,614	0,098
	3	2,600	**35,617**	0,612	0,027	0,012	**57,749**	1,146	0,082
	4	0,065	0,189	0,007	0,000	0,001	0,158	0,004	0,000
	5	0,059	0,037	0,000	0,000	0,009	0,082	0,001	0,000
	6	0,073	0,025	0,000	0,000	0,016	0,076	0,000	0,000
	7	0,059	0,019	0,001	0,000	0,030	0,070	0,000	0,000
	8	0,053	0,011	0,000	0,000	0,044	0,055	0,001	0,000
SC(p)	1	0,000	8,344	6,609	0,511	0,000	3,964	1,385	0,093
	2	0,003	61,966	2,279	0,105	0,000	55,156	2,776	0,169
	3	2,042	**17,485**	0,313	0,015	0,012	**35,283**	0,728	0,044
	4	0,049	0,045	0,000	0,000	0,001	0,083	0,002	0,000
	5	0,071	0,025	0,000	0,000	0,007	0,076	0,001	0,000
	6	0,057	0,016	0,000	0,000	0,013	0,063	0,000	0,000
	7	0,036	0,009	0,000	0,000	0,025	0,056	0,000	0,000
	8	0,017	0,003	0,000	0,000	0,027	0,035	0,001	0,000

The numbers represent the percentage the model selection criterion has in choosing the cell corresponding to the lag and number of cointegration vectors in 100,000 runs.
The true lag-cointegrating vectors are identified by boldface numbers.

Table A.2 Performance of the IC(p, s) information criterion in selecting p and s

Johansen tested coint.vectors(r)		0			**1**			2			3		
	Selected rank(s)	1	2	3	**1**	2	3	1	2	3	1	2	3
	Selected lag(p)												
Sample size=100													
	1	-	-	-	-	-	-	-	-	-	-	-	-
	2	0.002	0,000	0,000	39,049	0,001	0,000	1,218	0,000	0,000	0,056	0,000	0,000
	3	0,301	0,000	0,000	**56,341**	0,003	0,000	1,559	0,000	0,000	0,053	0,000	0,000
AIC(p, s)	4	0,004	0,000	0,000	1,186	0,001	0,000	0,070	0,000	0,000	0,001	0,000	0,000
	5	0,000	0,000	0,000	0,114	0,001	0,000	0,012	0,000	0,000	0,000	0,000	0,000
	6	0,000	0,000	0,000	0,020	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000
	7	0,000	0,000	0,000	0,006	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000
	8	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	1	-	-	-	-	-	-	-	-	-	-	-	-
	2	0,002	0,000	0,000	55,563	0,000	0,000	1,888	0,000	0,000	0,081	0,000	0,000
	3	0,267	0,000	0,000	**40,855**	0,000	0,000	1,207	0,000	0,000	0,043	0,000	0,000
HQ(p, s)	4	0,000	0,000	0,000	0,088	0,000	0,000	0,005	0,000	0,000	0,000	0,000	0,000
	5	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	6	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	7	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	8	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	1	-	-	-	-	-	-	-	-	-	-	-	-
	2	0,004	0,000	0,000	70,971	0,000	0,000	2,574	0,000	0,000	0,113	0,000	0,000
	3	0,221	0,000	0,000	**25,204**	0,000	0,000	0,887	0,000	0,000	0,025	0,000	0,000
SC($p.s$)	4	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	5	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	6	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	7	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	8	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Sample size=200													
	1	-	-	-	-	-	-	-	-	-	-	-	-
	2	0,000	0,000	0,000	18,797	0,000	0,000	0,681	0,000	0,000	0,038	0,000	0,000
	3	0,000	0,000	0,000	**77,065**	0,002	0,000	2,260	0,000	0,000	0,145	0,000	0,000
AIC(p, s)	4	0,000	0,000	0,000	0,908	0,000	0,000	0,035	0,000	0,000	0,001	0,000	0,000
	5	0,000	0,000	0,000	0,063	0,000	0,000	0,002	0,000	0,000	0,000	0,000	0,000
	6	0,000	0,000	0,000	0,003	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	7	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	8	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	1	-	-	-	-	-	-	-	-	-	-	-	-
	2	0,000	0,000	0,000	33,952	0,000	0,000	1,370	0,000	0,000	0,086	0,000	0,000
	3	0,000	0,000	0,000	**62,576**	0,000	0,000	1,877	0,000	0,000	0,111	0,000	0,000
HQ(p, s)	4	0,000	0,000	0,000	0,027	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000
	5	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	6	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	7	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	8	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	1	-	-	-	-	-	-	-	-	-	-	-	-
	2	0,000	0,000	0,000	50,983	0,000	0,000	2,351	0,000	0,000	0,146	0,000	0,000
	3	0,000	0,000	0,000	**45,028**	0,000	0,000	1,416	0,000	0,000	0,076	0,000	0,000
SC(p, s)	4	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	5	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	6	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	7	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	8	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000

The numbers represent the percentage the IC(p,s) simultaneous model selection criterion has in choosing the cell corresponding to the lag rank and number of cointegrating vectors in 100,000 runs. The true lag rank cointegration vectors are identified by boldface numbers and the best lag rank cointegration vectors chosen by the criteria are underlined.

Appendix B: VAR Restrictions for the DGPs

Consider the vector autoregressive, VAR(3), model:

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + A_3 y_{t-3} + \varepsilon_t \tag{16}$$

with parameters: $A_1 = \begin{bmatrix} a_{11}^1 & a_{12}^1 & a_{12}^1 \\ a_{21}^1 & a_{22}^1 & a_{22}^1 \\ a_{31}^1 & a_{32}^1 & a_{32}^1 \end{bmatrix}$, $A_2 = \begin{bmatrix} a_{11}^2 & a_{12}^2 & a_{12}^2 \\ a_{21}^2 & a_{22}^2 & a_{22}^2 \\ a_{31}^2 & a_{32}^2 & a_{32}^2 \end{bmatrix}$ and $A_3 = \begin{bmatrix} a_{11}^3 & a_{12}^3 & a_{12}^3 \\ a_{21}^3 & a_{22}^3 & a_{22}^3 \\ a_{31}^3 & a_{32}^3 & a_{32}^3 \end{bmatrix}$. We consider the cointegration vectors $\beta = \begin{bmatrix} \beta_{11} \\ \beta_{21} \\ \beta_{31} \end{bmatrix}$, the cofeature vectors $\tilde{\beta} = \begin{bmatrix} \tilde{\beta}_{11} & \tilde{\beta}_{12} \\ \tilde{\beta}_{21} & \tilde{\beta}_{22} \\ \tilde{\beta}_{31} & \tilde{\beta}_{32} \end{bmatrix}$ and the adjustment matrix $\alpha = \begin{bmatrix} \alpha_{11} \\ \alpha_{21} \\ \alpha_{31} \end{bmatrix}$. The long-run relation is defined by $\alpha\beta' = (A_1 + A_2 + A_3 - I_3)$. Thus, the VECM representation is:

$$\Delta y_t = \alpha\beta' y_{t-1} - (A_2 + A_3)\Delta y_{t-1} - A_3 \Delta y_{t-2} + \varepsilon_t \tag{17}$$

We can rewrite Equation (17) as a VAR(1):

$$\xi_t = F\ \xi_{t-1} + v_t \tag{18}$$

where $\xi_t = \begin{bmatrix} \triangle y_t \\ \triangle y_{t-1} \\ \beta' y_t \end{bmatrix}$, $F = \begin{bmatrix} -(A_2 + A_3) & -A_3 & \alpha \\ I_3 & 0 & 0 \\ -\beta(A_2 + A_3) & -\beta' A_3 & \beta'\alpha + 1 \end{bmatrix}$ and $v_t = \begin{bmatrix} \varepsilon_t \\ 0 \\ \beta'\varepsilon_t \end{bmatrix}$

1) Short-run restrictions (WF)

We now impose the common cyclical restrictions (i) and (ii) on model (16). Let, $G = -[R_{21}K + R_{31}]$, $K = [(R_{32} - R_{31})/(R_{21} - R_{22})]$, $R_{j1} = \tilde{\beta}_{j1}/\tilde{\beta}_{11}$, $R_{j2} = \tilde{\beta}_{j2}/\tilde{\beta}_{12}$ $(j = 2, 3)$ and $S = \beta_{11}G + \beta_{21}K + \beta_{31}$

(*i*) $\tilde{\beta}' A_3 = 0 => \ A_3 = \begin{bmatrix} -Ga_{31}^3 & -Ga_{32}^3 & -Ga_{33}^3 \\ -Ka_{31}^3 & -Ka_{32}^3 & -Ka_{33}^3 \\ -a_{31}^3 & -a_{32}^3 & -a_{33}^3 \end{bmatrix}$

(*ii*) $\tilde{\beta}'(A_2 + A_3) = 0 => \tilde{\beta}' A_2 = 0 => \ A_2 = \begin{bmatrix} -Ga_{31}^2 & -Ga_{32}^2 & -Ga_{33}^2 \\ -Ka_{31}^2 & -Ka_{32}^2 & -Ka_{33}^2 \\ -a_{31}^2 & -a_{32}^2 & -a_{33}^2 \end{bmatrix}$

2) Long-run restrictions (cointegration)

The cointegration restrictions are specified by (iv) and (v):

(iv) $\beta'(A_2 + A_3) = [-(a_{31}^2 + a_{31}^3)S \quad -(a_{32}^2 + a_{32}^3)S \quad -(a_{33}^2 + a_{33}^3)S]$ and $\beta' A_3 = [-a_{31}^3 S \quad -a_{32}^3 S \quad -a_{33}^3 S]$

(v) $\beta'\alpha + 1 = \beta = \begin{bmatrix} \beta_{11} & \beta_{21} & \beta_{31} \end{bmatrix} \begin{bmatrix} \alpha_{11} \\ \alpha_{21} \\ \alpha_{31} \end{bmatrix} + 1 = \beta_{11}\alpha_{11} + \beta_{21}\alpha_{21} + \beta_{31}\alpha_{31} + 1$

Taking into account the short- and long-run restrictions, the companion matrix F can be represented as:

$$F = \begin{bmatrix} -(A_2 + A_3) & -A_3 & \alpha \\ I_3 & 0 & 0 \\ -\beta(A_2 + A_3) & -\beta' A_3 & \beta'\alpha + 1 \end{bmatrix}$$

$$= \begin{bmatrix} -G(a_{31}^2 + a_{31}^3) & -G(a_{32}^2 + a_{32}^3) & -G(a_{33}^2 + a_{33}^3) & -Ga_{31}^3 & -Ga_{32}^3 & -Ga_{33}^3 & \alpha_{11} \\ -K(a_{31}^2 + a_{31}^3) & -G(a_{32}^2 + a_{32}^3) & -G(a_{33}^2 + a_{33}^3) & -Ka_{31}^3 & -Ka_{32}^3 & -Ka_{33}^3 & \alpha_{21} \\ -(a_{31}^2 + a_{31}^3) & -G(a_{32}^2 + a_{32}^3) & -(a_{33}^2 + a_{33}^3) & -a_{31}^3 & -a_{32}^3 & -a_{33}^3 & \alpha_{31} \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ -(a_{31}^2 + a_{31}^3)S & -(a_{32}^2 + a_{32}^3)S & -(a_{33}^2 + a_{33}^3)S & -a_{31}^3 S & -a_{32}^3 S & -a_{33}^3 S & b \end{bmatrix}$$

with $b = \beta'\alpha + 1 = \beta_{11}\alpha_{11} + \beta_{21}\alpha_{21} + \beta_{31}\alpha_{31} + 1$

3) Covariance stationary restrictions

Equation (18) will be covariance stationary if all eigenvalues of matrix F lie inside the unit circle. That is, eigenvalue of matrix F is a number λ such that:

$$|F - \lambda I_7| = 0 \tag{19}$$

The solution of (19) is:

$$\lambda^7 + \Omega\lambda^6 + \Theta\lambda^5 + \Psi\lambda^4 = 0 \tag{20}$$

where the parameters Ω, Θ, and Ψ are: $\Omega = G(a_{31}^2 + a_{31}^3) + K(a_{32}^2 + a_{32}^3) + a_{33}^2 + a_{33}^3 - b$, $\Theta = Ga_{31}^3 + Ka_{32}^3 - (a_{33}^2 + a_{33}^3)b - Gb(a_{31}^2 + a_{31}^3) - Kb(a_{32}^2 + a_{32}^3) + \alpha_{31}S(a_{33}^2 + a_{33}^3) + S\alpha_{21}(a_{32}^2 + a_{32}^3) + S\alpha_{11}(a_{31}^2 + a_{31}^3) + a_{33}^3$ and $\Psi = -a_{33}^3 b - Ga_{31}^3 b - Ka_{32}^3 b + \alpha_{31}a_{33}^3 S + a_{32}^3 S\alpha_{21} + a_{31}^3 S\alpha_{11}$, and the first four roots are $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$. We calculated the parameters of matrices A_1, A_2 and A_3 as functions of roots (λ_5, λ_6 and λ_7) and free parameters. Hence, we have three roots satisfying Equation (20)

$$\lambda^3 + \Omega\lambda^2 + \Theta\lambda + \Psi = 0 \tag{21}$$

for λ_5, we have: $\lambda_5^3 + \Omega\lambda_5^2 + \Theta\lambda_5 + \Psi = 0$ ....$Eq1$

for λ_6, we have: $\lambda_6^3 + \Omega\lambda_6^2 + \Theta\lambda_6 + \Psi = 0$ ....$Eq2$

for λ_7, we have: $\lambda_7^3 + \Omega\lambda_7^2 + \Theta\lambda_7 + \Psi = 0$ ....$Eq3$

Solving Equations 1, 2 and 3 yields: $\Omega = -\lambda_7 - \lambda_6 - \lambda_5$, $\Theta = \lambda_6\lambda_7 + \lambda_6\lambda_5 + \lambda_5\lambda_7$ and $\Psi = -\lambda_5\lambda_6\lambda_7$. Equating these parameters with the relations above we have:

$$
\begin{aligned}
a_{31}^2 &= -(-Ka_{32}^2 - Ka_{32}^2 b + \alpha_{31} S a_{33}^2 - \lambda^6\lambda^7 - \lambda^6 - \lambda^7 - a_{33}^2 b - \lambda^5\lambda^6\lambda^7 + b \\
&- \lambda^5\lambda^7 - \lambda^5\lambda^6 - a_{33}^2 + Sa_{32}^2\alpha_{21} - \lambda^5)/(S\alpha_{11} - G - Gb)
\end{aligned}
$$

$$
\begin{aligned}
a_{32}^3 &= (-S^2\lambda^7\alpha_{11}\alpha_{31} - b^2\lambda^7 G - \lambda^6 G b^2 + b\lambda^7 S\alpha_{11} + \lambda^6 S\alpha_{11} b - a_{31}^3 S\alpha_{11} G \\
&+ a_{31}^3 S^2 \alpha_{11}^2 - G a_{31}^3 b S\alpha_{11} - \lambda^5 G b^2 + \lambda^5 S\alpha_{11} b - \lambda^7\lambda^6\alpha_{31} SG - \lambda^7\lambda^5\alpha_{31} SG \\
&- S^2\alpha_{11}\lambda^5\alpha_{31} - S^2\alpha_{11}\lambda^6\alpha_{31} + S\lambda^5 G b\alpha_{31} + S\alpha_{31}\lambda^6 G b - \lambda^5\lambda^7\lambda^6 G \\
&+ \lambda^6\lambda^7 G b + \lambda^5\lambda^7 G b + \lambda^5\lambda^6 G b - S G b^2 \alpha_{31} + S^2\alpha_{11} b\alpha_{31} \\
&- S^2\alpha_{11}\alpha_{31} a_{33}^2 + S^2\alpha_{31}^2 a_{33}^2 G + S G^2 a_{31}^3 \alpha_{31} + S\alpha_{11} a_{33}^2 b + G b^3 - S\alpha_{11} b^2 \\
&- S^2\alpha_{11} K a_{32}^2 \alpha_{31} - S^2\alpha_{11}\alpha_{31} G a_{31}^3 + S^2 a_{32}^2 \alpha_{21} G\alpha_{31} - S a_{32}^2 \alpha_{21} G b \\
&+ S\alpha_{31} G^2 a_{31}^3 b - S\alpha_{31} a_{33}^2 G b + S\alpha_{11} K a_{32}^2 b + S\lambda^7 G b\alpha_{31} \\
&- \lambda^5\lambda^6\alpha_{31} SG - \lambda^5\lambda^7\lambda^6\alpha_{31} SG + \lambda^5\lambda^7\lambda^6 S\alpha_{11})/(S\alpha_{11} K\alpha_{31} - KG\alpha_{31} \\
&+ bG\alpha_{21} - K\alpha_{31} G b - S\alpha_{11}\alpha_{21} + G\alpha_{21})/S
\end{aligned}
$$

$$
\begin{aligned}
a_{33}^3 &= -(Kb^3G - \lambda^5 Gb^2K + S\alpha_{11}\lambda^6 K\lambda^7\lambda^5 \\
&+ Kb\lambda^7 S\alpha_{11} - Kb^2\lambda^7 G - S^2\alpha_{21}\lambda^7\alpha_{11} + \lambda^6 GbS\alpha_{21} \\
&+ S\alpha_{21}\lambda^7 Gb - \lambda^6 Gb^2K + \lambda^6 S\alpha_{11}Kb \\
&- \lambda^6 S^2\alpha_{11}\alpha_{21} + \lambda^5 GbS\alpha_{21} \\
&+ \lambda^5 S\alpha_{11}Kb - \lambda^5 S^2\alpha_{11}\alpha_{21} \\
&- \lambda^7\lambda^6 S\alpha_{21}G + Kb\lambda^7\lambda^6 G + Kb\lambda^7\lambda^5 G \\
&+ Kb\lambda^5\lambda^6 G - \lambda^7\lambda^6 KG\lambda^5 \\
&- S^2\alpha_{11}\alpha_{21}Ka_{32}^2 + S^2\alpha_{11}\alpha_{21}b \\
&- S^2\alpha_{11}\alpha_{21}a_{33}^2 + S^2\alpha_{21}^2 a_{32}^2 G \\
&- S\alpha_{11}Kb^2 + S\alpha_{21}G^2 a_{31}^3 - S\alpha_{21}Gb^2 \\
&+ S^2 a_{31}^3 K\alpha_{11}^2 - S^2\alpha_{11}\alpha_{21}Ga_{31}^3 \\
&+ S^2\alpha_{21}a_{33}^2 G\alpha_{31} + S\alpha_{11}K\hat{\ }2ba_{32}^2 \\
&+ S\alpha_{11}Kba_{33}^2 - S\alpha_{11}a_{31}^3 KG - S\alpha_{11}KbGa_{31}^3 \\
&- SKba_{33}^2 G\alpha_{31} + S\alpha_{21}G^2 a_{31}^3 b - S\alpha_{21}\lambda^5\lambda^6 G \\
&- S\alpha_{21}\lambda^5\lambda^7\lambda^6 G - S\alpha_{21}Ka_{32}^2 Gb \\
&- S\alpha_{21}\lambda^7\lambda^5 G)/(S\alpha_{11}K\alpha_{31} - KG\alpha_{31} + bG\alpha_{21} \\
&- K\alpha_{31}Gb - S\alpha_{11}\alpha_{21} + G\alpha_{21})/S
\end{aligned}
$$

We can calculate a_{31}^2, a_{32}^3 and a_{33}^3 fixing the set $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$. The values of a_{31}^3, a_{32}^2, a_{33}^2, λ_5, λ_6 and λ_7 are sorted independently from uniform distributions $(-0.9; 0.9)$. Hence, each parameter of the matrices A_1, A_2 and A_3 are defined and we can generate the DGPs of VAR(3) model with cointegration and WF restrictions.